



PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
Tormo *et al.*

Serial No.: Not Assigned

Filed: Concurrently

For: INHIBITION OF Bcl-2 PROTEIN
EXPRESSION BY LIPOSOMAL
ANTISENSE
OLIGODEOXYNUCLEOTIDES

Group Art Unit: Not Assigned

Examiner: Not Assigned

Atty. Dkt. No.: UTXC:504/WIM

EXPRESS MAIL MAILING LABEL
NUMBER TB 900063978 US
DATE OF DEPOSIT October 4, 1996
I hereby certify that this paper or fee is being deposited with the United States Postal Service "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 C.F.R. 1.10 on the date indicated above and is addressed to: Assistant Commissioner for Patents, Washington D.C. 20231.
Justin Whitley

STATEMENT AS REQUIRED UNDER 37 C.F.R. § 1.821(f)

ATTN: PATENT APPLICATIONS

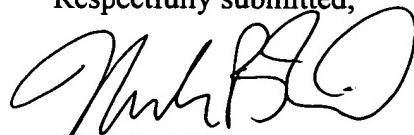
Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Submitted herewith is a computer readable form of the sequence listing and a paper copy of the sequence listing of those sequences in the captioned patent application. The computer readable form of the sequence listing is the same as the paper copy of the sequence listing. The

sequence information provided in the Specification is also the same as the sequence listing of the enclosed computer readable and paper forms of the sequence.

Respectfully submitted,



Mark B. Wilson
Reg. No. 37,259
Attorney for Applicants

ARNOLD, WHITE & DURKEE
P.O. Box 4433
Houston, Texas 77210-4433
(512) 418-3000

Date: October 4, 1996



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Tormo, Mar
Tari, Ana M.
Lopez-Berestein, Gabriel
- (ii) TITLE OF INVENTION: INHIBITION OF Bcl-2 PROTEIN EXPRESSION BY LIPOSOMAL ANTISENSE OLIGODEOXYNUCLEOTIDES
- (iii) NUMBER OF SEQUENCES: 7
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Arnold, White & Durkee
(B) STREET: P.O. Box 4433
(C) CITY: Houston
(D) STATE: Texas
(E) COUNTRY: United States of America
(F) ZIP: 77210
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US Unknown
(B) FILING DATE: Concurrently Herewith
(C) CLASSIFICATION: Unknown
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Wilson, Mark B.
(B) REGISTRATION NUMBER: 37,259
(C) REFERENCE/DOCKET NUMBER: UTXC:504
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (512) 418-3000
(B) TELEFAX: (512) 474-7577

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAGCGTGCGC CATCCTTC

18

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACGGTCCGCC ACTCCTTCCC

20

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTGAAGGGCT TCTTCC

16

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 5086 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1459..2175

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGCCCGGCC	CTCCGCGCCG	CCTGCCCGCC	CGCCCGCCGC	GCTCCCGCCC	GCCGCTCTCC	60
GTGGCCCCGC	CGCGCTGCCG	CCGCCGCCGC	TGCCAGCGAA	GGTGCCGGGG	CTCCGGGGCCC	120
TCCCTGCCGG	CGGCCGTCAG	CGCTCGGAGC	GAACTGCGCG	ACGGGAGGTC	CGGGAGGCGA	180
CCGTAGTCGC	GCCGCCGCGC	AGGACCAGGA	GGAGGAGAAA	GGGTGCGCAG	CCCGGAGGCG	240
GGGTGCGCCG	GTGGGGTGCA	GCGGAAGAGG	GGGTCCAGGG	GGGAGAACTT	CGTAGCAGTC	300
ATCCTTTTA	GGAAAAGAGG	GAAAAAATAA	AACCCTCCCC	CACCACCTCC	TTCTCCCCAC	360
CCCTCGCCGC	ACCACACACA	GCGCGGGCTT	CTAGCGCTCG	GCACCGGGCG	GCCAGGCGCG	420
TCCTGCCTTC	ATTATCCAG	CAGCTTTCG	GAAAATGCAT	TTGCTGTTCG	GAGTTAAC	480
AGAAGACGAT	TCCTGCCTCC	GTCCCCGGCT	CCTTCATCGT	CCCATCTCCC	CTGTCTCTCT	540
CCTGGGGAGG	CGTGAAGCGG	TCCC GTGGAT	AGAGATTCA	GCCTGTGTCC	GCGCGTGTGT	600
GCGCGCGTAT	AAATTGCCGA	GAAGGGAAA	ACATCACAGG	ACTTCTGCGA	ATACCGGACT	660
GAAAATTGTA	ATTCATCTGC	CGCCGCCGCT	GCCAAAAAAA	AACTCGAGCT	CTTGAGATCT	720

.CCGGTTGGGA TTCTCGCGA TTGACATTC TGTGAAGCAG AAGTCTGGGA ATCGATCTGG	780
AAATCCTCCT AATTTTACT CCCTCTCCCC CCGACTCCTG ATTCAATTGGG AAGTTTCAA	840
TCAGCTATAA CTGGAGAGTG CTGAAGATTG ATGGGATCGT TGCCTTATGC ATTTGTTTG	900
GTTTACAAA AAGGAAACTT GACAGAGGAT CATGCTGTAC TTAAAAAATA CAAGTAAGTC	960
TCGCACAGGA AATTGGTTA ATGTAACCTT CAATGGAAAC CTTTGAGATT TTTTACTTAA	1020
AGTGCATTCG AGTAAATTAA ATTTCCAGGC AGCTTAATAC ATTGTTTTA GCCGTGTTAC	1080
TTGTAGTGTG TATGCCCTGC TTTCACTCAG TGTGTACAGG GAAACGCACC TGATTTTTA	1140
CTTATTAGTT TGTTTTCT TTAACCTTTC AGCATCACAG AGGAAGTAGA CTGATATTAA	1200
CAATACTTAC TAATAATAAC GTGCCTCATG AAATAAAGAT CCGAAAGGAA TTGGAATAAA	1260
AATTCCTGC GTCTCATGCC AAGAGGGAAA CACCAGAAC AAGTGTCCG CGTGATTGAA	1320
GACACCCCT CGTCCAAGAA TGCAAAGCAC ATCCAATAAA ATAGCTGGAT TATAACTCCT	1380
CTTCTTCTC TGGGGGCCGT GGGGTGGGAG CTGGGGCGAG AGGTGCCGTT GGCCCCCGTT	1440
GCTTTCCCTC TGGGAAGG ATG GCG CAC GCT GGG AGA ACG GGG TAC GAC AAC	1491
Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn	
1 5 10	
 CGG GAG ATA GTG ATG AAG TAC ATC CAT TAT AAG CTG TCG CAG AGG GGC	1539
Arg Glu Ile Val Met Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly	
15 20 25	
 TAC GAG TGG GAT GCG GGA GAT GTG GGC GCC GCG CCC CCG GGG GCC GCC	1587
Tyr Glu Trp Asp Ala Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala	
30 35 40	
 CCC GCA CCG GGC ATC TTC TCC TCC CAG CCC GGG CAC ACG CCC CAT CCA	1635
Pro Ala Pro Gly Ile Phe Ser Ser Gln Pro Gly His Thr Pro His Pro	
45 50 55	
 GCC GCA TCC CGC GAC CCG GTC GCC AGG ACC TCG CCG CTG CAG ACC CCG	1683
Ala Ala Ser Arg Asp Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro	
60 65 70 75	
 GCT GCC CCC GGC GCC GCG GGG CCT GCG CTC AGC CCG GTG CCA CCT	1731
Ala Ala Pro Gly Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro	
80 85 90	
 GTG GTC CAC CTG GCC CTC CGC CAA GCC GGC GAC GAC TTC TCC CGC CGC	1779
Val Val His Leu Ala Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg	
95 100 105	
 TAC CGC GGC GAC TTC GCC GAG ATG TCC AGC CAG CTG CAC CTG ACG CCC	1827
Tyr Arg Gly Asp Phe Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro	
110 115 120	
 TTC ACC GCG CGG GGA CGC TTT GCC ACG GTG GTG GAG GAG CTC TTC AGG	1875
Phe Thr Ala Arg Gly Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg	
125 130 135	

GAC GGG GTG AAC TGG GGG AGG ATT GTG GCC TTC TTT GAG TTC GGT GGG Asp Gly Val Asn Trp Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly 140 145 150 155	1923
GTC ATG TGT GTG GAG AGC GTC AAC CGG GAG ATG TCG CCC CTG GTG GAC Val Met Cys Val Glu Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp 160 165 170	1971
AAC ATC GCC CTG TGG ATG ACT GAG TAC CTG AAC CGG CAC CTG CAC ACC Asn Ile Ala Leu Trp Met Thr Glu Tyr Leu Asn Arg His Leu His Thr 175 180 185	2019
TGG ATC CAG GAT AAC GGA GGC TGG GAT GCC TTT GTG GAA CTG TAC GGC Trp Ile Gln Asp Asn Gly Gly Trp Asp Ala Phe Val Glu Leu Tyr Gly 190 195 200	2067
CCC AGC ATG CGG CCT CTG TTT GAT TTC TCC TGG CTG TCT CTG AAG ACT Pro Ser Met Arg Pro Leu Phe Asp Phe Ser Trp Leu Ser Leu Lys Thr 205 210 215	2115
CTG CTC AGT TTG GCC CTG GTG GGA GCT TGC ATC ACC CTG GGT GCC TAT Leu Leu Ser Leu Ala Leu Val Gly Ala Cys Ile Thr Leu Gly Ala Tyr 220 225 230 235	2163
CTG AGC CAC AAG TGAAGTCAAC ATGCCTGCC CAAACAAATA TGCAAAAGGT Leu Ser His Lys	2215
TCACTAAAGC AGTAGAAATA ATATGCATTG TCAGTGATGT ACCATGAAAC AAAGCTGCAG	2275
GCTGTTAACG AAAAAATAAC ACACATATAA ACATCACACA CACAGACAGA CACACACACA	2335
CACAACAATT AACAGTCTTC AGGCAAAACG TCGAATCAGC TATTTACTGC CAAAGGGAAA	2395
TATCATTAT TTTTACATT ATTAAGAAAA AAGATTATT TATTTAAGAC AGTCCCATCA	2455
AAACTCCGTC TTTGGAAATC CGACCACTAA TTGCCAAACA CCGCTTCGTG TGGCTCCACC	2515
TGGATGTTCT GTGCCTGTAA ACATAGATTG GCTTCCATG TTGTTGGCCG GATCACCATC	2575
TGAAGAGCAG ACGGATGGAA AAAGGACCTG ATCATTGGGG AAGCTGGCTT TCTGGCTGCT	2635
GGAGGCTGGG GAGAAGGTGT TCATTCACCTT GCATTTCTTT GCCCTGGGG CGTGATATTA	2695
ACAGAGGGAG GGTTCCCGTG GGGGAAAGTC CATGCCTCCC TGGCCTGAAG AAGAGACTCT	2755
TTGCATATGA CTCACATGAT GCATACCTGG TGGGAGGAAA AGAGTTGGGA ACTTCAGATG	2815
GACCTAGTAC CCACTGAGAT TTCCACGCCG AAGGACAGCG ATGGGAAAAA TGCCCTTAAA	2875
TCATAGGAAA GTATTTTTT AAGCTACCAA TTGTGCCGAG AAAAGCATT TAGCAATTAA	2935
TACAATATCA TCCAGTACCT TAAACCTGA TTGTGTATAT TCATATATT TGGATACGCA	2995
CCCCCCAACT CCCAATACTG GCTCTGTCTG AGTAAGAAC AGAACCTCT GGAACCTTGAG	3055
GAAGTGAACA TTTCGGTGAC TTCCGATCAG GAAGGCTAGA GTTACCCAGA GCATCAGGCC	3115
GCCACAAGTG CCTGCTTTA GGAGACCGAA GTCCGCAGAA CCTACCTGTG TCCCAGCTTG	3175

GAGGCCTGGT CCTGGAAC TG AGCCGGGCC TCACTGGCCT CCTCCAGGG A TGATCACAG	3235
GGTAGTGTGG TCTCCGAATG TCTGGAAAGCT GATGGATGGA GCTCAGAATT C CACTGTCAA	3295
GAAAGAGCAG TAGAGGGGTG TGGCTGGGCC TGTCACCCCTG GGGCCCTCCA GGTAGGCCCG	3355
TTTCACGTG GAGCATAGGA GCCACGACCC TTCTTAAGAC ATGTATCACT GTAGAGGGAA	3415
GGAACAGAGG CCCTGGGCCT TCCTATCAGA AGGACATGGT GAAGGCTGGG AACGTGAGGA	3475
GAGGCAATGG CCACGGCCA TTTGGCTGT AGCACATGGC ACGTTGGCTG TGTGGCCTTG	3535
GCCACCTGTG AGTTAAAGC AAGGCTTAA ATGACTTGG AGAGGGTCAC AAATCCTAAA	3595
AGAACGATTG AAGTGAGGTG TCATGGATT ATTGACCCCT GTCTATGGAA TTACATGTAA	3655
AACATTATCT TGTCACTGTA GTTGGTTTT ATTTGAAAAC CTGACAAAAA AAAAGTTCCA	3715
GGTGTGGAAT ATGGGGTTA TCTGTACATC CTGGGGCATT AAAAAAAAAT CAATGGTGGG	3775
GAACATATAAA GAAGTAACAA AAGAAGTGAC ATCTTCAGCA AATAAAACTAG GAAATTTTT	3835
TTTCTTCCAG TTTAGAATCA GCCTGAAAC ATTGATGGAA TAACTCTGTG GCATTATTGC	3895
ATTATATACC ATTTATCTGT ATTAACCTTG GAATGTACTC TGTTCAATGT TTAATGCTGT	3955
GGTTGATATT TCGAAAGCTG CTTTAAAAAA ATACATGCAT CTCAGCGTTT TTTGTTTT	4015
AATTGTATTT AGTTATGGCC TATACACTAT TTGTGAGCAA AGGTGATCGT TTTCTGTTT	4075
AGATTTTAT CTCTTGATTC TTCAAAAGCA TTCTGAGAAG GTGAGATAAG CCCTGAGTCT	4135
CAGCTACCTA AGAAAAACCT GGATGTCACT GGCCACTGAG GAGCTTGTT TCAACCAAGT	4195
CATGTGCATT TCCACGTCAA CAGAATTGTT TATTGTGACA GTTATATCTG TTGTCCTTT	4255
GACCTTGTGTT CTTGAAGGTT TCCTCGTCCC TGGGCAATTG CGCATTAAAT TCATGGTATT	4315
CAGGATTACA TGCATGTTTG GTTAAACCCA TGAGATTCA TCAGTTAAA ATCCAGATGG	4375
CGAATGACCA GCAGATTCAA ATCTATGGTG GTTGACCTT TAGAGAGTTG CTTTACGTGG	4435
CCTGTTCAA CACAGACCC CCCAGAGCCC TCCTGCCCTC CTTCCGCGGG GGCTTCTCA	4495
TGGCTGTCCT TCAGGGTCTT CCTGAAATGC AGTGGTCGTT ACGCTCCACC AAGAAAGCAG	4555
GAAACCTGTG GTATGAAGCC AGACCTCCCC GGCGGCCCTC AGGGAACAGA ATGATCAGAC	4615
CTTGTGAAATGA TTCTAATTT TAAGCAAAAT ATTATTTAT GAAAGGTTA CATTGTCAA	4675
GTGATGAATA TGGAATATCC AATCCTGTGC TGCTATCCTG CCAAAATCAT TTTAATGGAG	4735
TCAGTTGCA GTATGCTCCA CGTGGTAAGA TCCTCCAAGC TGCTTAGAA GTAACAATGA	4795
AGAACGTGGA CGTTTTAAT ATAAAGCCTG TTTGTCTTT TGTTGTTGTT CAAACGGGAT	4855
TCACAGAGTA TTTGAAAAAT GTATATATAT TAAGAGGTCA CGGGGGCTAA TTGCTAGCTG	4915
GCTGCCTTT GCTGTGGGT TTTGTTACCT GGTTTAATA ACAGTAAATG TGCCCAGCCT	4975

CTTGGCCCCA GAACTGTACA GTATTGTGGC TGCAC TTGCT CTAAGAGTAG TTGATGTTGC 5035
ATTTTCCTTA TTGTTAAAAA CATGTTAGAA GCAATGAATG TATATAAAAG C 5086

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 239 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
1 5 10 15

Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
20 25 30

Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
35 40 45

Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp
50 55 60

Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
65 70 75 80

Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala
85 90 95

Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe
100 105 110

Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
115 120 125

Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp
130 135 140

Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
145 150 155 160

Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
165 170 175

Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
180 185 190

Gly Gly Trp Asp Ala Phe Val Glu Leu Tyr Gly Pro Ser Met Arg Pro
195 200 205

Leu Phe Asp Phe Ser Trp Leu Ser Leu Lys Thr Leu Leu Ser Leu Ala
210 215 220

Leu Val Gly Ala Cys Ile Thr Leu Gly Ala Tyr Leu Ser His Lys

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 911 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 147..761

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TGATTGAAGA CACCCCTCG TCCAAGAACAT CAAAGCACAT CCAATAAAAT AGCTGGATTA	60
TAACTCCTCT TCTTTCTCTG GGGGCCGTGG GGTGGGAGCT GGGGCGAGAG GTGCCGTTGG	120
CCCCCGTTGC TTTTCCTCTG GGAAGG ATG GCG CAC GCT GGG AGA ACG GGG TAC Met Ala His Ala Gly Arg Thr Gly Tyr	173
1 5	
GAC AAC CGG GAG ATA GTG ATG AAG TAC ATC CAT TAT AAG CTG TCG CAG Asp Asn Arg Glu Ile Val Met Lys Tyr Ile His Tyr Lys Leu Ser Gln	221
10 15 20 25	
AGG GGC TAC GAG TGG GAT GCG GGA GAT GTG GGC GCC GCG CCC CCG GGG Arg Gly Tyr Glu Trp Asp Ala Gly Asp Val Gly Ala Ala Pro Pro Gly	269
30 35 40	
GCC GCC CCC GCA CCG GGC ATC TTC TCC TCC CAG CCC GGG CAC ACG CCC Ala Ala Pro Ala Pro Gly Ile Phe Ser Ser Gln Pro Gly His Thr Pro	317
45 50 55	
CAT CCA GCC GCA TCC CGC GAC CCG GTC GCC AGG ACC TCG CCG CTG CAG His Pro Ala Ala Ser Arg Asp Pro Val Ala Arg Thr Ser Pro Leu Gln	365
60 65 70	
ACC CCG GCT GCC CCC GGC GCC GCG GGG CCT GCG CTC AGC CCG GTG Thr Pro Ala Ala Pro Gly Ala Ala Gly Pro Ala Leu Ser Pro Val	413
75 80 85	
CCA CCT GTG GTC CAC CTG GCC CTC CGC CAA GCC GGC GAC GAC TTC TCC Pro Pro Val Val His Leu Ala Leu Arg Gln Ala Gly Asp Asp Phe Ser	461
90 95 100 105	
CGC CGC TAC CGC GGC GAC TTC GCC GAG ATG TCC AGC CAG CTG CAC CTG Arg Arg Tyr Arg Gly Asp Phe Ala Glu Met Ser Ser Gln Leu His Leu	509
110 115 120	
ACG CCC TTC ACC GCG CGG GGA CGC TTT GCC ACG GTG GTG GAG GAG CTC Thr Pro Phe Thr Ala Arg Gly Arg Phe Ala Thr Val Val Glu Glu Leu	557
125 130 135	
TTC AGG GAC GGG GTG AAC TGG GGG AGG ATT GTG GCC TTC TTT GAG TTC Phe Arg Asp Gly Val Asn Trp Gly Arg Ile Val Ala Phe Phe Glu Phe	605
140 145 150	

.GGT GGG GTC ATG TGT GTG GAG AGC GTC AAC CGG GAG ATG TCG CCC CTG Gly Gly Val Met Cys Val Glu Ser Val Asn Arg Glu Met Ser Pro Leu 155 160 165	653
GTG GAC AAC ATC GCC CTG TGG ATG ACT GAG TAC CTG AAC CGG CAC CTG Val Asp Asn Ile Ala Leu Trp Met Thr Glu Tyr Leu Asn Arg His Leu 170 175 180 185	701
CAC ACC TGG ATC CAG GAT AAC GGA GGC TGG GTA GGT GCA TCT GGT GAT His Thr Trp Ile Gln Asp Asn Gly Gly Trp Val Gly Ala Ser Gly Asp 190 195 200	749
GTG AGT CTG GGC TGAGGCCACA GGTCCGAGAT CGGGGGTTGG AGTGCGGGTG Val Ser Leu Gly 205	801
GGCTCCTGGG CAATGGGAGG CTGTGGAGCC GGCGAAATAA AATCAGAGTT GTTGCTTCCC	861
GGCGTGTCCC TACCTCCTCC TCTGGACAAA GCGTTCACTC CCAACCTGAC	911

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 205 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met 1 5 10 15
Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala 20 25 30
Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile 35 40 45
Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp 50 55 60
Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala 65 70 75 80
Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala 85 90 95
Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe 100 105 110
Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly 115 120 125
Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp 130 135 140

Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
145 150 155 160

Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
165 170 175

Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
180 185 190

Gly Gly Trp Val Gly Ala Ser Gly Asp Val Ser Leu Gly
195 200 205